



Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers

JARED HAMWOOD, DAVID ALONSO-CANEIRO,* SCOTT A. READ, STEPHEN J. VINCENT, AND MICHAEL J. COLLINS

Contact Lens and Visual Optics Laboratory, School of Optometry and Vision Science, Queensland University of Technology, Brisbane, Queensland, Australia

*d.alonsocaneiro@qut.edu.au

Abstract: Deep learning strategies, particularly convolutional neural networks (CNNs), are especially suited to finding patterns in images and using those patterns for image classification. The method is normally applied to an image patch and assigns a class weight to the patch; this method has recently been used to detect the probability of retinal boundary locations in OCT images, which is subsequently used to segment the OCT image using a graph-search approach. This paper examines the effects of a number of modifications to the CNN architecture with the aim of optimizing retinal layer segmentation, specifically the effect of patch size as well as the network architecture design on CNN performance and subsequent layer segmentation. The results demonstrate that increasing patch size can improve the performance of the classification and provides a more reliable segmentation in the analysis of retinal layer characteristics in OCT imaging. Similarly, this work shows that changing aspects of the CNN network design can also significantly improve the segmentation results. This work also demonstrates that the performance of the method can change depending on the number of classes (i.e. boundaries) used to train the CNN, with fewer classes showing an inferior performance due to the presence of similar image features between classes that can trigger false positives. Changes in the network (patch size and or architecture) can be applied to provide a superior segmentation performance, which is robust to the class effect. The findings from this work may inform future CNN development in OCT retinal image analysis.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

OCIS codes: (110.4500) Optical coherence tomography; (100.0100) Image processing; (100.2960) Image analysis.

References and links

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and et, "Optical coherence tomography," *Science* **254**(5035), 1178–1181 (1991).
2. J. F. de Boer, R. Leitgeb, and M. Wojtkowski, "Twenty-five years of optical coherence tomography: the paradigm shift in sensitivity and speed provided by Fourier domain OCT," *Biomed. Opt. Express* **8**(7), 3248–3280 (2017).
3. S. A. Read, M. J. Collins, S. J. Vincent, and D. Alonso-Caneiro, "Macular retinal layer thickness in childhood," *Retina* **35**(6), 1223–1233 (2015).
4. P. Jin, H. Zou, J. Zhu, X. Xu, J. Jin, T. C. Chang, L. Lu, H. Yuan, S. Sun, B. Yan, J. He, M. Wang, and X. He, "Choroidal and retinal thickness in children with different refractive status measured by swept-source optical coherence tomography," *Am. J. Ophthalmol.* **168**, 164–176 (2016).
5. S. A. Read, D. Alonso-Caneiro, and S. J. Vincent, "Longitudinal changes in macular retinal layer thickness in pediatric populations: Myopic vs non-myopic eyes," *PLoS One* **12**(6), e0180462 (2017).
6. I. I. Bussell, G. Wollstein, and J. S. Schuman, "OCT for glaucoma diagnosis, screening and detection of glaucoma progression," *Br. J. Ophthalmol.* **98**(2 Suppl 2), ii15–ii19 (2014).
7. A. E. Fung, G. A. Lalwani, P. J. Rosenfeld, S. R. Dubovy, S. Michels, W. J. Feuer, C. A. Puliafito, J. L. Davis, H. W. Flynn, Jr., and M. Esquiabro, "An optical coherence tomography-guided, variable dosing regimen with intravitreal ranibizumab (Lucentis) for neovascular age-related macular degeneration," *Am. J. Ophthalmol.* **143**(4), 566–583 (2007).

8. W. J. Lee, Y. K. Kim, Y. W. Kim, J. W. Jeoung, S. H. Kim, J. W. Heo, H. G. Yu, and K. H. Park, "Rate of macular ganglion cell-inner plexiform layer thinning in glaucomatous eyes with vascular endothelial growth factor inhibition," *J. Glaucoma* **26**(11), 980–986 (2017).
9. G. Virgili, F. Menchini, V. Murro, E. Peluso, F. Rosa, and G. Casazza, "Optical coherence tomography (OCT) for detection of macular oedema in patients with diabetic retinopathy," *Cochrane Database Syst. Rev.* **7**(7), CD008081 (2011).
10. D. Alonso-Caneiro, S. A. Read, S. J. Vincent, M. J. Collins, and M. Wojtkowski, "Tissue thickness calculation in ocular optical coherence tomography," *Biomed. Opt. Express* **7**(2), 629–645 (2016).
11. A. Baghaie, Z. Yu, and R. M. D'Souza, "State-of-the-art in retinal optical coherence tomography image analysis," *Quant. Imaging Med. Surg.* **5**(4), 603–617 (2015).
12. D. C. DeBuc, "A review of algorithms for segmentation of retinal image data using optical coherence tomography," in *Image Segmentation*, P.-G. Ho, ed. (InTech, 2011), pp. 15–54.3.
13. H. Ishikawa, D. M. Stein, G. Wollstein, S. Beaton, J. G. Fujimoto, and J. S. Schuman, "Macular segmentation with optical coherence tomography," *Invest. Ophthalmol. Vis. Sci.* **46**(6), 2012–2017 (2005).
14. D. Koozekanani, K. Boyer, and C. Roberts, "Retinal thickness measurements from optical coherence tomography using a Markov boundary model," *IEEE Trans. Med. Imaging* **20**(9), 900–916 (2001).
15. V. Kajić, B. Považay, B. Hermann, B. Hofer, D. Marshall, P. L. Rosin, and W. Drexler, "Robust segmentation of intraretinal layers in the normal human fovea using a novel statistical model based on texture and shape analysis," *Opt. Express* **18**(14), 14730–14744 (2010).
16. S. J. Chiu, X. T. Li, P. Nicholas, C. A. Toth, J. A. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," *Opt. Express* **18**(18), 19413–19428 (2010).
17. M. K. Garvin, M. D. Abramoff, X. Wu, S. R. Russell, T. L. Burns, and M. Sonka, "Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," *IEEE Trans. Med. Imaging* **28**(9), 1436–1447 (2009).
18. A. Mishra, A. Wong, K. Bizheva, and D. A. Clausi, "Intra-retinal layer segmentation in optical coherence tomography images," *Opt. Express* **17**(26), 23719–23728 (2009).
19. P. P. Srinivasan, L. A. Kim, P. S. Mettu, S. W. Cousins, G. M. Comer, J. A. Izatt, and S. Farsiu, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomed. Opt. Express* **5**(10), 3568–3577 (2014).
20. B. Keller, D. Cunefer, D. S. Grewal, T. H. Mahmoud, J. A. Izatt, and S. Farsiu, "Length-adaptive graph search for automatic segmentation of pathological features in optical coherence tomography images," *J. Biomed. Opt.* **21**(7), 076015 (2016).
21. S. Sankar, D. Sidibé, Y. Cheung, T. Wong, E. Lamoureux, D. Milea, and F. Mériaudeau, "Classification of SD-OCT volumes for DME detection: an anomaly detection approach," in *Medical Imaging 2016: Computer-Aided Diagnosis*, (International Society for Optics and Photonics, 978520 (2016).
22. B. Hassan, G. Raja, T. Hassan, and M. Usman Akram, "Structure tensor based automated detection of macular edema and central serous retinopathy using optical coherence tomography images," *J. Opt. Soc. Am. A* **33**(4), 455–463 (2016).
23. D. Sidibé, S. Sankar, G. Lemaître, M. Rastgoo, J. Massich, C. Y. Cheung, G. S. W. Tan, D. Milea, E. Lamoureux, T. Y. Wong, and F. Mériaudeau, "An anomaly detection approach for the identification of DME patients using spectral domain optical coherence tomography images," *Comput. Methods Programs Biomed.* **139**, 109–117 (2017).
24. M. W. M. Wintergerst, T. Schultz, J. Birtel, A. K. Schuster, N. Pfeiffer, S. Schmitz-Valckenberg, F. G. Holz, and R. P. Finger, "Algorithms for the automated analysis of age-related macular degeneration biomarkers on optical coherence tomography: a systematic review," *Transl. Vis. Sci. Technol.* **6**(4), 10 (2017).
25. J. Massich, M. Rastgoo, G. Lemaître, C. Y. Cheung, T. Y. Wong, D. Sidibé, and F. Mériaudeau, "Classifying DME vs normal SD-OCT volumes: a review," in *Pattern Recognition (ICPR)*, 2016 23rd International Conference on, (IEEE, 2016), 1297–1302.
26. B. Liefers, F. G. Venhuizen, V. Schreur, B. van Ginneken, C. Hoyng, S. Fauser, T. Theelen, and C. I. Sánchez, "Automatic detection of the foveal center in optical coherence tomography," *Biomed. Opt. Express* **8**(11), 5160–5178 (2017).
27. Y. Xu, K. Yan, J. Kim, X. Wang, C. Li, L. Su, S. Yu, X. Xu, and D. D. Feng, "Dual-stage deep learning framework for pigment epithelium detachment segmentation in polypoidal choroidal vasculopathy," *Biomed. Opt. Express* **8**(9), 4061–4076 (2017).
28. K. A. Vermeer, J. van der Schoot, H. G. Lemij, and J. F. de Boer, "Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images," *Biomed. Opt. Express* **2**(6), 1743–1756 (2011).
29. A. Lang, A. Carass, M. Hauser, E. S. Sotirchos, P. A. Calabresi, H. S. Ying, and J. L. Prince, "Retinal layer segmentation of macular OCT images using boundary classification," *Biomed. Opt. Express* **4**(7), 1133–1152 (2013).
30. A. Ben-Cohen, D. Mark, I. Kovler, D. Zur, A. Barak, M. Iglicki, and R. Soferman, "Retinal layers segmentation using fully convolutional network in OCT images," *RSIP Vision* 2017.
31. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-assisted Intervention* (Springer, 2015), 234–241.

32. F. G. Venhuizen, B. van Ginneken, B. Liefers, M. J. J. P. van Grinsven, S. Fauser, C. Hoyng, T. Theelen, and C. I. Sánchez, "Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks," *Biomed. Opt. Express* **8**(7), 3292–3316 (2017).
33. A. G. Roy, S. Conjeti, S. P. K. Karri, D. Sheet, A. Katouzian, C. Wachinger, and N. Navab, "ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks," *Biomed. Opt. Express* **8**(8), 3627–3642 (2017).
34. L. Fang, D. Cuneffare, C. Wang, R. H. Guymier, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express* **8**(5), 2732–2744 (2017).
35. S. A. Read, D. Alonso-Caneiro, S. J. Vincent, and M. J. Collins, "Longitudinal changes in choroidal thickness and eye growth in childhood," *Invest. Ophthalmol. Vis. Sci.* **56**(5), 3103–3112 (2015).
36. S. A. Read, M. J. Collins, and S. J. Vincent, "Light exposure and eye growth in childhood," *Invest. Ophthalmol. Vis. Sci.* **56**(11), 6779–6787 (2015).
37. R. F. Spaide, H. Koizumi, and M. C. Pozzoni, "Enhanced depth imaging spectral-domain optical coherence tomography," *Am. J. Ophthalmol.* **146**(4), 496–500 (2008).
38. S. Y. Park, S. M. Kim, Y.-M. Song, J. Sung, and D.-I. Ham, "Retinal thickness and volume measured with enhanced depth imaging optical coherence tomography," *Am. J. Ophthalmol.* **156**(3), 557–566 (2013).
39. A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*, (ACM, 2015), 689–692.
40. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Adv. Neural Inf. Process. Syst.* **1**, 1097–1105 (2012).
41. H. R. Roth, L. Lu, J. Liu, J. Yao, A. Seff, K. Cherry, L. Kim, and R. M. Summers, "Improving computer-aided detection using convolutional neural networks and random view aggregation," *IEEE Trans. Med. Imaging* **35**(5), 1170–1181 (2016).
42. D. Cuneffare, L. Fang, R. F. Cooper, A. Dubra, J. Carroll, and S. Farsiu, "Open source software for automatic detection of cone photoreceptors in adaptive optics ophthalmoscopy using convolutional neural networks," *Sci. Rep.* **7**(1), 6620 (2017).
43. H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 132–139 (2016).
44. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431–3440 (2015).
45. V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 807–814 (2010).
46. L. N. Smith and N. Topin, "Deep convolutional neural network design patterns," *arXiv preprint arXiv:1611.00847* (2016).
47. S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747* (2016).
48. E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.* **1**(1), 269–271 (1959).
49. H. Li, R. Zhao, and X. Wang, "Highly efficient forward and backward propagation of convolutional neural networks for pixelwise classification," *arXiv preprint arXiv:1412.4526* (2014).
50. C. Farabet, C. Couprie, L. Najman, and Y. Lecun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(8), 1915–1929 (2013).
51. P. H. Pinheiro and R. Collobert, "Recurrent convolutional neural networks for scene labeling," in *31st International Conference on Machine Learning (ICML) (No. EPFL-CONF-199822)*, (2014).
52. M. Bashkansky and J. Reintjes, "Statistics and reduction of speckle in optical coherence tomography," *Opt. Lett.* **25**(8), 545–547 (2000).
53. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J. Biomed. Opt.* **4**(1), 95–105 (1999).

1. Introduction

Optical coherence tomography (OCT) has transformed the imaging of the eye, providing high-resolution cross-sectional images of the ocular tissue that are now commonly used in clinical practice and research to better understand the eye in both health and disease [1, 2]. Although qualitative assessment of OCT images can be used to inform clinical decision making, including the classification and detection of eye diseases, quantification of the acquired images is necessary to better understand the eye's normal development and the impact of common eye conditions such as myopia upon eye morphology [3–5], and to more reliably facilitate early disease detection, the tracking of disease progression and monitoring of treatment efficacy [6–9]. Estimates of tissue thickness, derived from quantitative OCT image analysis remains the most common way to characterize the B-scan images acquired from the instrument [10]. Thus, layer segmentation is a fundamental task in OCT image

analysis, and the development of reliable OCT image segmentation methods has been the focus of numerous previous studies.

A number of papers have reviewed the rich literature involving OCT image analysis, particularly layer segmentation methods [11-12], which includes a large variety of approaches such as analysis of intensity variation and adaptive thresholding [13], intensity-based Markov boundary models [14], texture and shape analysis [15], graph theory techniques [16], multi-surface graph cuts approach [17], and active contour segmentation models [18]. These techniques are in general a set of ad hoc rules applied to the image to extract the boundaries of interest. Although the methods have shown significant merit, they do not always generalize well and changes in the image may mean that changes in the set of rules are also required. Machine learning methods develop their own rules based on the provided training data (images plus boundary position) and may be more resilient to variations in the data that are commonly encountered in clinical images. Thus in recent years, machine learning methods have emerged as a useful tool for a range of retinal OCT image analysis applications, including the classification of OCT images into different disease groups and for the detection of pathological features in images [19–23]. Reviews of this research area can be found elsewhere [24, 25]. Other applications include, automatic detection of the foveal center in patients with age-related macular degeneration [26] and the detection and segmentation of retinal pigment epithelium detachments [27].

For layer segmentation, a number of studies have used machine learning methods to detect the boundaries between retinal layers. Vermeer et al. [28] first extracted several features from individual A-scans, then used these features in a Support Vector Machine to classify each pixel within the A-scan to segment six retinal layers. Features extracted included the intensity, neighboring intensity to various lengths, and neighboring gradients to various lengths. Lang et al. [29] used a Random Forest Classifier (RFC) to classify features around a pixel into nine layer classes, then either a Canny edge detector or graph-search was used to segment the layers into boundaries. Ben-Cohen et al. [30] utilized a combination of the U-net fully convolutional architecture [31], Sobel edge detection, and graph-search to identify four retinal boundaries. U-net first identified the full retinal layers, then the Sobel edge detector was used to segment these layers, and finally Dijkstra's graph-search algorithm was used to predict the boundaries between layers. Venhuizen et al. [32] also used a U-net based network to calculate the total thickness of the retina. U-net was used to classify parts of the scan as either retinal or another part of the eye, then this classification was thresholded to create a binary image of the retina. Roy et al. [33] created an encoder-decoder framework similar to U-net, called ReLayNet, to segment the image into retinal tissue (eight layers) and intra-layer fluid. ReLayNet was used as an end-to-end approach, with the output being a completely segmented image with no further processing necessary.

Recently, Fang et al. [34] proposed a convolutional neural network (CNN) and graph-search method (termed as CNN-GS) for the automatic segmentation of nine layer boundaries in OCT retinal images of patients with non-exudative age-related macular degeneration. The work follows Chiu's [16] original OCT segmentation model based on graph theory, but replaces the edge-detection step with a CNN. The CNN predicts the boundary location providing a per-layer probability image, which is then used to trace the boundary using graph-search methods. The method runs a patch (window) of fixed size through the image to provide the probability map for a particular boundary to be present in the center of that window. The authors adopted a previous CNN architecture, proposed for a different data set, for the classification of the patch in OCT retinal boundaries. Their model operates on a square input patch size of 33x33 pixels. In this paper, an in-depth analysis of the effect of using different patch sizes and CNN architectures on segmentation performance is presented. The aim of this work is to better understand the effect of patch size as well as network architecture on CNN performance and subsequent layer segmentation results, in order to improve the performance of the classification and optimize the outcomes from the OCT retinal layer

segmentation. The findings of this work may inform future CNN development in OCT imaging analysis using deep learning strategies.

2. Methods

2.1 Demographics and OCT data set

A retrospective data set of retinal OCT images was used for this study and a detailed description of the study participants and procedures have been provided in a number of previous publications [5, 35, 36]. Briefly, this retrospective data is from a longitudinal study examining macular retinal layer thickness in childhood involving 101 children (13.1 ± 1.4 years) with a range of refractive errors. Retinal OCT images were collected on each child at four study visits, conducted every 6 months over an 18-month period, although for this work, OCT images were analyzed from the first visit only. The study was approved by the Queensland University of Technology human research ethics committee and all study procedures followed the tenets of the Declaration of Helsinki. All participants enrolled had normal vision in both eyes, no history or evidence of ocular disease, injury or surgery and no manifest hyperopic refractive errors of greater than + 1.25 DS.

High-resolution cross-sectional retinal images were collected using the Heidelberg Spectralis (Heidelberg Engineering, Heidelberg, Germany) SD-OCT instrument. This device uses a super luminescent diode with a central wavelength of 870 nm for OCT scanning (capturing 40,000 A-scans per second), and provides cross-sectional retinal OCT images with an axial digital resolution of 3.9 μm . The images captured were 496 pixels deep and 1,536 pixels wide, with a total area of 761,856 pixels, with scale of 3.9 μm per pixel deep, and 5.7 μm per pixel wide.

At each study visit, the participants had 2 series of 6 high resolution foveal centered radial OCT scan lines acquired using the instrument's Enhanced Depth Imaging (EDI) mode (with each radial scan line separated by 30 degrees). The EDI mode is typically used to enhance the visibility of the choroid [37], and it has been shown that retinal thickness measures from EDI scans are comparable to those collected using the Spectralis instrument's conventional imaging mode [38]. To improve the image signal to noise ratio, 30 frames were averaged using the instrument's automatic real time eye tracking feature.

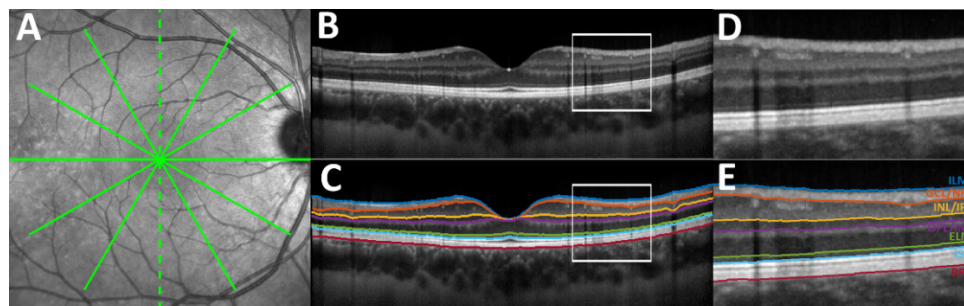


Fig. 1. Example of the spectral domain OCT B-scan captured using the instrument's high resolution scanning protocol including (A) the fundus image and (B) B-scan. The B-scan is shown with the seven boundaries of interest (C), including the retinal pigment epithelium (RPE), inner segment ellipsoid (ISe), external limiting membrane (ELM), outer plexiform layer/inner nuclear layer (OPL/INL) boundary, inner nuclear layer/inner plexiform layer (INL/IPL) boundary, ganglion cell layer/nerve fiber layer (GCL/NFL) boundary and the inner limiting membrane (ILM). Insets show a zoomed version of the B-scan (D) without and (E) with the boundaries of interest.

The exported OCT images were analyzed using custom written software. Initially, an automated graph based method [3, 16] was used to segment the boundaries of 7 different retinal layers, including: the outer boundary of the retinal pigment epithelium (RPE), the inner boundary of the inner segment ellipsoid zone (ISe), the inner boundary of the external

limiting membrane (ELM), the boundary between the outer plexiform layer and inner nuclear layer (OPL/INL), the boundary between the inner nuclear layer and the inner plexiform layer (INL/IPL), the boundary between the ganglion cell layer and the nerve fiber layer (GCL/NFL) and the inner boundary of the inner limiting membrane (ILM) (Fig. 1). An experienced observer, masked to the demographic and refractive details of the participants, then checked the integrity of the automated segmentation of each boundary and manually corrected any segmentation errors.

2.2 Overview of image processing methods

The method used for the segmentation of OCT images follows a similar procedure to that used by Fang et al [34]. This is a two stage method consisting of an initial CNN followed by a graph-search procedure. The CNN computes a probability map that indicates the likelihood of a retinal boundary lying at any given pixel, and this probability map is then used as an input to the graph-search method that traces from the top left of the image to the bottom right to calculate the most likely path of the retinal boundary of interest. All training and testing of networks was computed on an Nvidia Titan Xp using MATLAB r2017a and VLfeat's MatConvNet library [39].

The Fang method uses a well-known CNN, the CIFAR-CNN architecture [40, 41], which was originally designed and tested to classify small 32x32 color images into 10 classes, and the authors adopted this network for retinal OCT image layer segmentation. In this paper, we explore the proposed method and examine the effects of a number of modifications to the CNN architecture with the aim of optimizing retinal layer segmentation. The modifications take into account the distribution of the information (retinal boundaries within the tissue) within the OCT image and adapts the patch size to better capture the richness of the information and improve image segmentation. Figure 2 provides an overview of the proposed methods, which have been divided into two major steps. The first step involves the training of the CNN network, using a data set with seven labelled retinal boundaries. The second step involves the testing of the trained network performance to produce probability maps and then subsequent graph-search to trace these boundary locations.

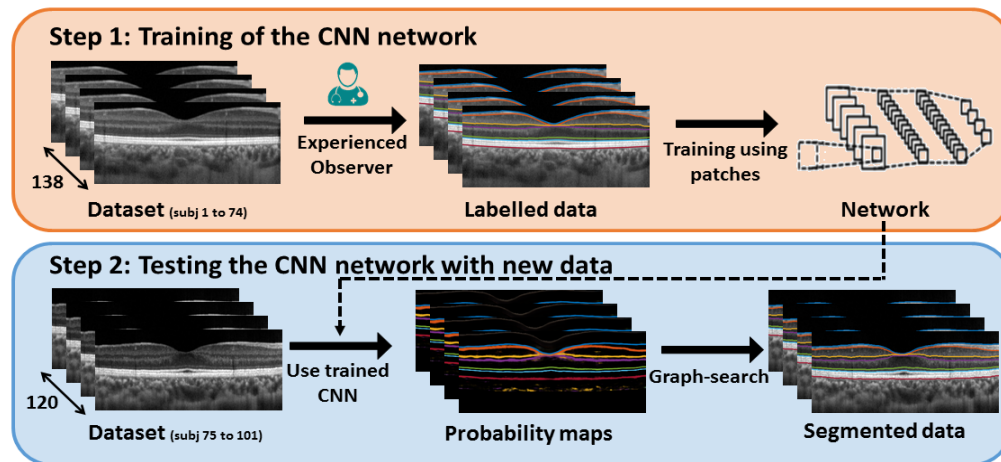


Fig. 2. Overview of the proposed method with the two major steps involved in the process. In the training step (top section) the different proposed networks can be substituted to evaluate the effect of different networks on the performance of the segmentation task.

2.3 Convolutional neural networks

Convolutional Neural Networks (CNNs) are classifiers often used for image analysis [41]. In ophthalmic applications, CNNs have been used for a number of different purposes, including

cone detection in high-resolution retinal images obtained using adaptive optics [42], and retinal vessel segmentation [43]. CNNs consist of many layers, normally arranged into blocks which find local features, activate based on these findings, and sub-sample to create a smaller but more feature-rich input for the next block [44]. The most commonly used layers for these blocks are combinations of convolutional layers, activation layers and sub-sampling layers. The convolutional layers take an $X1 \times Y1 \times Z1$ sized input and convolute this to a $1 \times 1 \times Z2$ output. The activation layers are normally implemented using Rectified Linear Units (ReLU) [45] which activate when the input is greater than an arbitrary threshold. The sub-sampling layers, normally average or max pooling layers, reduce the size of the input while preserving the information. Other layers that are used less commonly, but still appear in most networks, are fully connected layers, which are simply convolutional layers connecting to all values in the input, and softmax layers, which are used to calculate the final probabilities of each class.

In general, a network block consists of a convolutional layer, a ReLU layer, and a sub-sampling layer. Each block reduces the height and breadth of the input, but increases the depth of the input, until the final classification is performed [46]. CNNs are normally trained by a gradient descent method over multiple iterations, where each sample is used multiple times. This gradient descent algorithm seeks to minimize the error with respect to various weights throughout the network. Most modern networks can consist of millions of parameters which govern the convolutional or activation functions within the network. In most cases the data set used for training is too large to fit within the memory of a single GPU, so training is often done in several smaller batches (a stochastic gradient descent [47]).

2.4 Patch prediction and boundary segmentation

The first stage in the CNN-GS method is the prediction of the boundary locations. Fang et al. [34] pre-processed images by applying an intensity normalization, for completeness this step was still included in the procedure. The normalization procedure (originally presented in [28]) ensures the data is located between a 0 to 1 range and removes outliers by looking for maximum values of a smooth filter image. To train the network, small patches of arbitrary dimensions were sampled once per column (A-scan) of the OCT image, centered in the boundary, with a randomly chosen non-boundary sample patch to provide a null or background sample. To preserve the weighting of the network, any column without all layers present was omitted from the training or testing sets. Each layer was given an arbitrary class label from 1 to 8, including the background class.

A total of 138 OCT images from 70 subjects were used to form the training set, with 28 of these scans used as validation samples during training. Training was done by stochastic gradient descent with momentum [47], in batches of 1024 randomly chosen samples at a time. Given the scans in this project were collected in a radial pattern, an equal number of scans from each orientation were used during training and validation. For the prediction of the boundaries, a CNN was trained to predict the probability of a sample patch belonging to any of the seven boundary classes, or the background class in the case it did not belong to any boundary. The probability map contains values between zero and one, where values close to one indicate a high likelihood of the boundary being present in that particular pixel position (Fig. 3).

To perform image segmentation from the probability maps, a graph-search method is used to find the highest weighted probability path from one side of the scan to the other side for each boundary. This follows a similar approach to Chiu et al [16]. Thus, the probability map represents a graph of nodes, where each pixel corresponds to a node. Dijkstra's algorithm [48] determines the preferred path between any two nodes on the entire graph by calculating the lowest weight between them. By initializing the start-node and end-node at each side of the probability map, the detected path matches with the layer of interest given the appropriate weight map. The weight assigned to the edges connecting adjacent nodes a and b is calculated as follows; $W_{ab} = 2 \cdot (g_a + g_b) + W_{min}$. Where g_a , g_b are the gradient information at nodes a and

b respectively with $g_a, g_b \in [0; 1]$ and W_{\min} is a constant value of 10^{-5} . Here a 4-node connection is used (right, top, left, and bottom). After the weights maps are calculated, Dijkstra's algorithm is used to determine the lowest weighted path of a graph between the start and end nodes. This process is repeated for all seven non-background classes (i.e. boundaries), until a prediction has been created for each layer.

2.5 Network design rationale

There were several networks created during testing, however only the best performing networks will be discussed in detail. The CIFAR network, used by Fang et al. [34], was used as the baseline for comparison purposes. This network originally used a 32x32 size input, but this was adapted to a 33x33 image patch to center the boundary at the center of the patch. The aim of this work is to examine the effects of a number of CNN modifications (patch size and network design) with the goal of optimizing retinal layer segmentation.

Patch size: When the original patch was used, some boundaries were easily confused with neighboring boundaries without the full topological frame. This negative effect was particularly significant when reducing the class size (see section 3.3). For this reason, we hypothesized that an increase in patch size may improve performance for these particular layers. Previous studies have shown that using a larger image patch in CNNs can lead to better classification accuracy, since CNNs can capture more contextual information to make the decision [49–51]. Thus, a larger network was tested (65x65 pixels patch size), which was designed to resemble the baseline network topology. However, it is important to note that increasing the patch size also requires changes in the network, since some of the layers (fully-connected layers) need to have a fixed size input by definition. Thus, the change of patch size also implies a change in the layer architecture.

Table 1. Architecture of the four different CNNs used in this work. For comparison, the baseline network (33x33 CIFAR) proposed by Fang et al. [34] was included. The CIFAR 65x65 CNN extends the original input size to take advantage of the distribution of the information within the OCT image. The complex CNN network explores the benefit of altering the network architecture. The number in round brackets “()” indicates the number of filters, number in angle brackets “<>” indicates stride, and number in curly brackets “{}” indicates padding. Fully Con indicates a Fully Connected layer. For the 4 class network (section 3.3) the final layer will be “1x1 Fully Connected (4) <1> {0}”.

CIFAR based CNN network		Complex CNN network	
33x33 greyscale patch	65x65 greyscale patch	33x33 greyscale patch	65x65 greyscale patch
5x5 Convolutional (32) <1> {2}		5x5 Convolutional (32) <1> {0}	
3x3 Max Pooling <2> {0 1 0 1}		ReLU	
ReLU		3x3 Convolutional (32) <1> {0}	
5x5 Convolutional (32) <1> {0}		ReLU	
ReLU		3x3 Convolutional (64) <1> {0}	
3x3 Avg. Pooling <2> {0 1 0 1}		2x2 Max Pooling <2> {1}	
5x5 Convolutional (64) <1> {0}		3x3 Convolutional (64) <1> {0}	
ReLU		ReLU	
3x3 Avg. Pooling <2> {0 1 0 1}		3x3 Convolutional (128) <1> {0}	
4x4 Fully Con (64) <1> {0}	8x8 Fully Con (64) <1> {0}	2x2 Max Pooling <2> {1}	
ReLU		13x13 Fully Con (128) <1> {0}	13x13 Fully Conn (128) <1> {0}
		1x1 Fully Con (8) <1> {0}	1x1 Fully Con (8) <1> {0}

Complex network design: The original CIFAR-10 network used in this work was design to handle a RGB object classification problem. Here, we proposed a more complex CNN network, by removing some of the pooling layers. Instead, convolutional layers can be used to reduce the size of the input if it is not zero padded. U-Net uses a similar principle to focus the result on a central area [31]. The benefit of having no zero padding is that there is no effect

from adding a large amount of 0's around the border of the image, and the output only depends on the actual data. For comparison purposes, the network was tested with both patch sizes (33x33 and 65x65) explored in the previous version.

Table 1 provides the detail for the tested architectures and the four tested CNNs provide various patch sizes and network complexities to gain a better understanding of the CNN performance. The Fang et al. [34] network (CIFAR 33x33) is used as a baseline for our work.

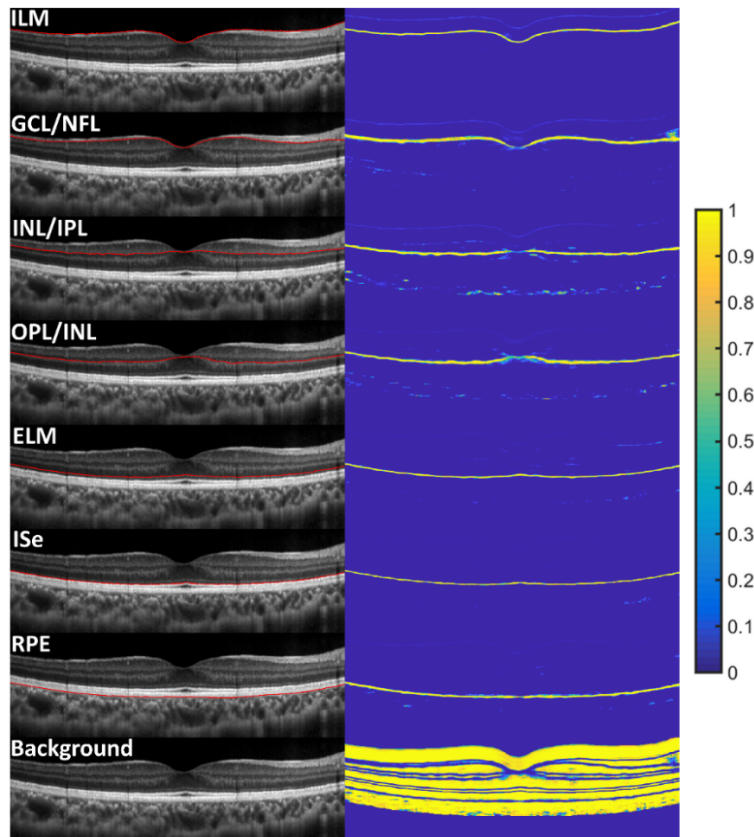


Fig. 3. Example OCT image with the layer of interest delineated in red (left) and the equivalent probability map for that layer (right), the last image displays the background (no boundary) class. For computational reasons the probability maps (values of one indicate high likelihood of the boundary to be present) are only computed a number of pixels equal to half the patch size above the ILM and below the RPE. The trained CNN CIFAR with a patch size of 65x65 was used to extract the seven probability maps associated with the boundaries of interest and the background class in this figure.

2.6 Evaluation

To evaluate the performance of the network, 120 OCT images from 20 subjects were used to create a test data set. All subjects in the test data set were different to those in the training set, including validation subjects. This ensured that the training data set was completely separate from the data set used for the testing. Similar to the training, an equal number of radial orientations were chosen. For all networks, the images were pre-processed with the same filter used by Fang et al. [34], but no additional pre-processing was performed in comparison to the baseline network.

When supplied with a test image, the image was first split into patches around each pixel. In the case of pixels where the patch fell outside the borders of the image, zeros were added to pad the patch. Once every pixel had its corresponding patch created, the patches were

provided to the network as inputs, and the resulting class probabilities became that pixel's values in the probability map. Once the probability of every pixel was calculated, the boundaries were segmented using the graph-search approach. These boundaries were then compared to the known truth values (from the original segmentation data [5] that were examined and manually corrected by an experienced trained human observer) to calculate the mean and absolute error and the standard deviation of the error. In a large portion of the scans used there was sub-optimal image quality and poor scan detail at the extreme edges of the images, as well as anatomical features where layers were not present (e.g. the optic nerve head), so for the calculation of error, the images were truncated by 100 pixels on either edge.

Repeatability was also measured by comparing the predicted thickness between each series of 2 repeated measurements for a single participant. Only 14 participants of the 20 used for testing had consecutive measurements present for all orientations in both series, so only 84 pairs of B-scans were used for assessing the repeatability.

3. Results

3.1 Overall performance

To compare the overall performance of the network, the difference between the predicted boundary location for each network architecture and the true boundary location was calculated using both mean error and mean absolute error on a column by column basis, with the standard deviation of both errors also included. All values are given in pixels throughout the paper.

Table 2. Difference in boundary position between the different network architectures and patch sizes and the manual observer for the entire data set (120 B-scans). The results are reported in mean values and (standard deviation) in pixel units.

CIFAR based CNN architecture				
	33x33 patch		65x65 patch	
Retinal boundary	Mean error	Mean abs. error	Mean error	Mean abs. error
ILM	0.05 (2.64)	0.85 (2.50)	-0.09 (1.47)	0.82 (1.23)
GCL/NFL	0.36 (7.74)	3.17 (7.06)	0.24 (3.94)	1.76 (3.53)
INL/IPL	-0.81 (3.94)	1.52 (3.72)	-0.24 (2.16)	1.26 (1.78)
OPL/IPL	-0.07 (3.33)	1.27 (3.08)	-0.47 (1.82)	1.30 (1.35)
ELM	-0.05 (1.37)	0.70 (1.17)	-0.13 (1.16)	0.68 (0.95)
ISe	-0.23 (1.87)	0.62 (1.78)	-0.28 (1.29)	0.64 (1.15)
RPE	0.10 (2.50)	0.98 (2.31)	-0.42 (3.64)	1.05 (3.51)

Complex CNN architecture				
	33x33 patch		65x65 patch	
Retinal boundary	Mean error	Mean abs. error	Mean error	Mean abs. error
ILM	-0.09 (0.96)	0.58 (0.77)	-0.10 (1.49)	0.59 (1.37)
GCL/NFL	1.17 (5.34)	1.92 (5.12)	0.06 (3.01)	1.15 (2.78)
INL/IPL	-0.06 (1.41)	0.68 (1.23)	-0.03 (1.35)	0.69 (1.17)
OPL/IPL	-0.01 (1.52)	0.79 (1.31)	-0.13 (1.35)	0.78 (1.10)
ELM	-0.07 (0.90)	0.49 (0.75)	-0.06 (1.13)	0.50 (1.01)
ISe	-0.08 (1.07)	0.49 (0.95)	-0.00 (1.03)	0.47 (0.91)
RPE	-0.08 (0.88)	0.57 (0.68)	-0.13 (1.43)	0.59 (1.31)

The mean absolute error is a measure of how far all points are from the truth-value regardless of whether these points fall above or below the truth. It can be observed that the mean absolute per-layer error for the 65x65 networks are comparable or slightly lower than

the 33x33 network (mean absolute error for all layers 1.30 pixels [33x33 CIFAR] and 1.07 pixels [65x65 CIFAR]), (Fig. 4 and Table 2 top). However, there is also a more significant decrease in the standard deviation of the large networks relative to the 33x33 networks. This represents both a more consistently correct segmentation outcome, and that the errors were generally clustered very close to the truth.

In comparison, the mean error is the average of both positive and negative errors, so this mean error value can be balanced out by similar amounts of positive and negative errors even if the predictions are not close to the truth-value. Across the different networks, the values of the mean error are small (all below 1 pixel), with only minimal differences across the different networks Table 2 (top). Similarly, the standard deviation shows small differences across most of the retinal boundaries for the larger patch image.

Table 2 (bottom) presents the effect of the complex networks for the same two patch sizes. There are significantly smaller errors in comparison to the CIFAR CNN. Similar to the CIFAR network, the mean absolute per-layer error for the larger 65x65 complex CNN networks are comparable or slightly lower than the 33x33 network (mean absolute error for all layer 0.78 pixels [33x33 Complex] and 0.71 pixels [65x65 Complex]).

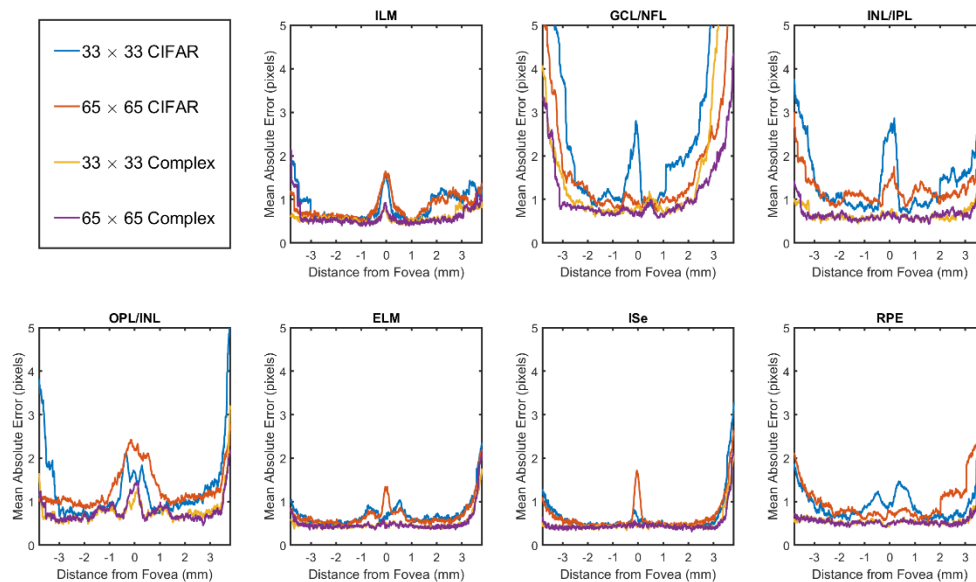


Fig. 4. Mean absolute error difference in boundary position between the different network architectures and the manual observer for the entire data set.

Figure 4 provides the comparison for each of the seven boundaries considered in the study, with the graphs providing a mean absolute error for the seven boundaries of interest for each of the four considered CNN architectures. When viewed on a layer by layer basis (Fig. 4), the performance of the methods varies between the different layers as well as with distance from the foveal center within each layer. Although the larger CIFAR network yields a lower mean absolute error in most boundaries in comparison to the smaller CIFAR network, the more complex networks appear to provide more significant improvements. Across the different networks, the layers that showed slightly worse performance were the NFL followed by the INL/IPL, particularly for the 33x33 network, but performance improved significantly using the larger patch size 65x65 or the Complex CNN.

Figure 5 provides an illustration of the boundary prediction from an A-scan of a representative B-Scan image, with three (33x33 CIFAR and both Complex networks) corresponding per-layer probability profiles for that A-scan (i.e. a cross-section of the per-layer probability maps). For this particular example, it can be observed how the different

architectures affects the outcomes. Both of the complex networks show narrower peaks in the probability maps, where the peak of the probability coincides well with the true location marked by the observer. The baseline network (CIFAR 33x33) provides a wider profile with less distinct peaks and more false positives (i.e. secondary peaks) from other layers. Overall, the large (65x65) Complex network appears to provide better results and shows a closer agreement with the experienced observer. Figure 6 presents the boundary comparison (large Complex network versus manual) for five B-scans from five different subjects, illustrating this close agreement.

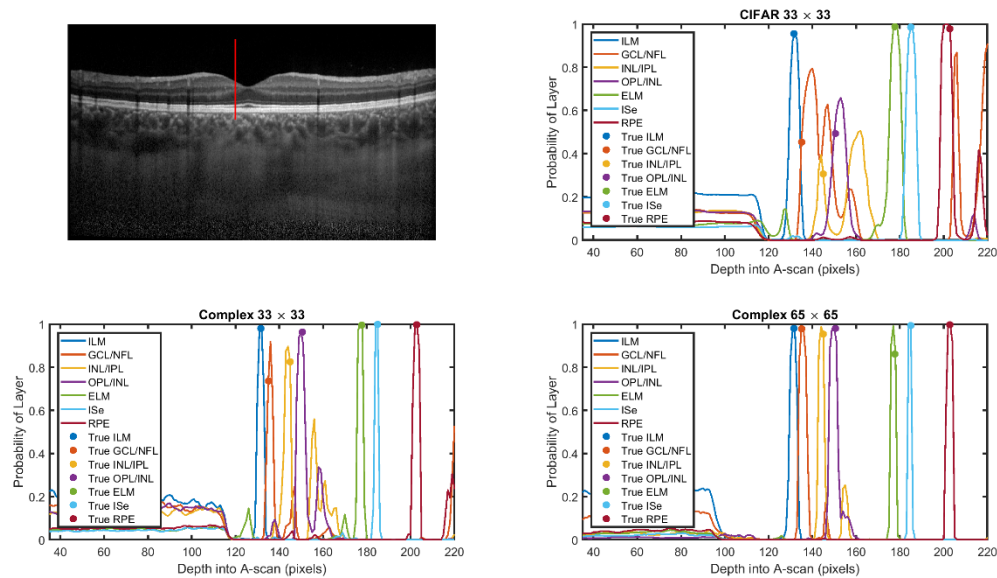


Fig. 5. Probability profile for the different considered networks for an individual A-scan in a representative OCT B-Scan (top left). Red line in OCT B-Scan represents the A-scan location. Solid circles indicate the true boundary position, whereas the lines show the probability of the boundary location predicted by the different CNN networks, including the CIFAR 33x33 (top right), the Complex 33x33 (bottom left) and the Complex 65x65 (bottom right).

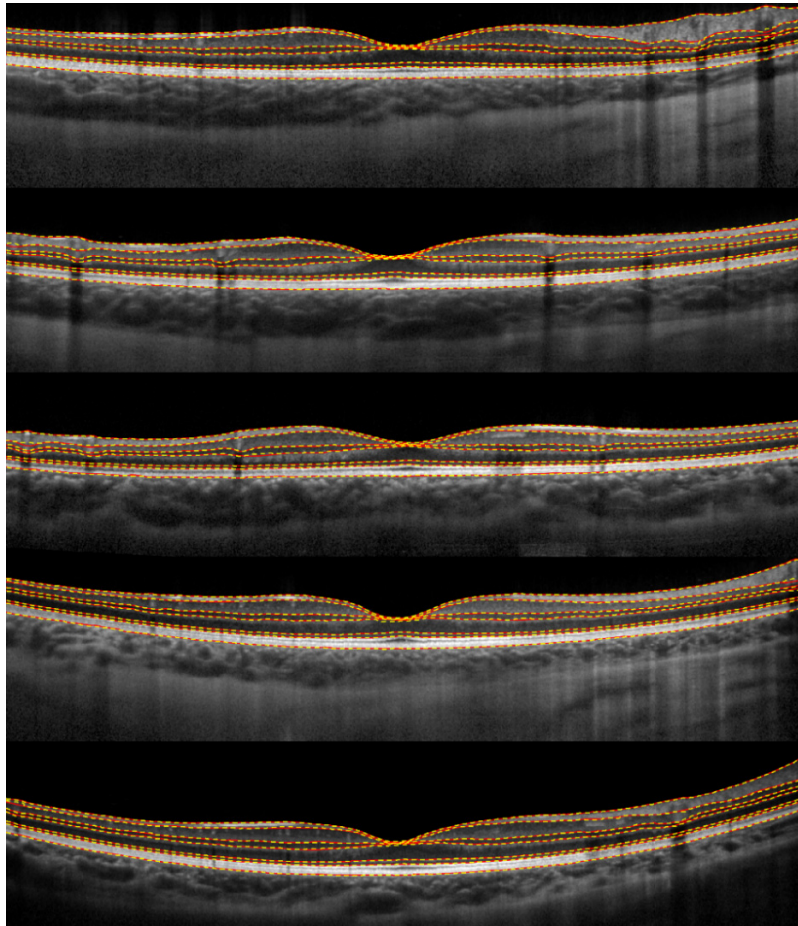


Fig. 6. Five B-scans from five different subjects are shown with the derived boundaries of interest. The B-scans represent a range of different examples (i.e. with typical variations in overall retinal thickness, thickness profiles and radial scan location). The segmentation based on the manual analysis (yellow) and the automatic analysis (red) is illustrated. The large Complex network (65x65) was used to generate the automatic results.

3.2 Repeatability results

Repeatability error was calculated for each retinal layer. The error was calculated by comparing the thickness from one series of measurements to the other series of measurements (i.e. a comparison between the two repeated series of 6 radial scans collected on each subject). Comparisons were only made between scans of the same orientation between series.

Table 3 presents both the mean thickness repeatability as well as the mean absolute thickness repeatability for baseline (33x33 CIFAR) and large complex (65x65 Complex) networks. The mean absolute thickness repeatability shows the larger differences between networks, with the 65x65 network exhibiting the best overall performance. The Complex large network shows the lowest errors across all boundaries, with a mean reduction in the mean absolute repeatability of 0.85 pixel [range 0.01 to 2.41 pixels] compared to the other network. For the mean absolute thickness repeatability, there is a larger margin in standard deviation, with the larger Complex 65x65 network showing substantially smaller standard deviations, indicative of more consistent performance, having smaller errors on average compared to the 33x33 CIFAR network.

Table 3. Thickness repeatability for the baseline (33x33 CIFAR) and large Complex (65x65 Complex) network architectures considered in the study, tested on a data set with 84 pairs of B-scans. The results are reported in mean values and (standard deviation) in pixel units. Thickness definition; total retinal (TR) thickness [RPE to ILM], the RPE + ISe thickness [RPE to ISe], the inner segment (IS) thickness [ISe to ELM] and the ONL + OPL thickness [ELM to OPL/INL], INL thickness [OPL/INL to INL/IPL], IPL + GCL thickness [INL/IPL to GCL/NFL] and NFL thickness [GCL/NFL to ILM].

	CIFAR 33x33 patch		Complex 65x65 patch	
Thickness	Mean repeatability	Mean abs. repeatability	Mean repeatability	Mean abs. repeatability
NFL	0.72 (10.66)	3.50 (10.09)	0.11 (3.50)	1.30 (3.25)
IPL + GCL	-0.72 (10.32)	3.89 (9.59)	0.09 (3.54)	1.48 (3.22)
INL	0.09 (4.76)	1.43 (4.54)	-0.08 (1.68)	0.87 (1.43)
ONL + OPL	0.00 (3.01)	1.02 (2.83)	0.08 (1.34)	0.81 (1.07)
IS	0.01 (1.24)	0.45 (1.15)	0.02 (0.89)	0.44 (0.77)
RPE + ISe	0.05 (1.48)	0.77 (1.26)	-0.02 (0.98)	0.53 (0.82)
TR	0.15 (3.73)	1.55 (3.39)	0.20 (2.11)	1.20 (1.75)

3.3 Effect of CNN classes on performance

To observe the effect that the number of classes has on the performance of the different networks, all four networks were trained again but using only 4 classes (background class plus three boundaries, including ILM, INL/IPL and RPE) and the results were compared with those results from the previous 8 classes (background class plus seven boundaries: ILM, GCL/NFL, INL/IPL, OPL/IPL, ELM, ISe and RPE). Table 4 provides the summary performance for the boundary position evaluated on the 120 OCT images from 20 subjects. The results demonstrate that the number of classes can have a negative effect on the performance of the network, since reducing the number of classes increases the boundary error. This increase is especially significant for the mean absolute error, but the standard deviation of the mean error also shows an increase.

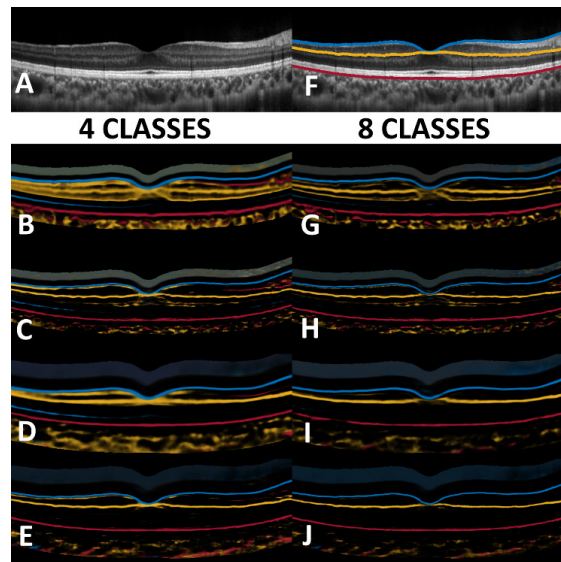


Fig. 7. Example B-scan (A) showing the probability maps for three layers of interest (F), (ILM in blue, INL/IPL in orange and RPE in red). Each color indicates a high probability of a boundary to be present in that location. The images show the effect on performance while training the CNN with different networks (B-G is the CIFAR CNN 33x33, C-H is the Complex CNN 33x33, D-I is the CIFAR CNN 65x65, and E-J is the Complex CNN 65x65) and different numbers of training classes (B-C-D-E 4 classes and G-H-I-J 8 classes).

This effect is likely linked to the higher number of false positive in the probability maps from the group with a lower number of classes, which will negatively impact the graph-search method. To show this, Fig. 7 compares the probability maps for a particular B-scan, only for the three layers of interest (ILM in blue, INL/IPL in orange and RPE in red). The color in the map indicates a high probability of a boundary to be present in that location. The image shows the effect on performance while training the CNN with different networks and training classes. Interestingly, for the larger and complex networks, the effect is reduced and the probability maps show a narrower path with fewer false positives, indicating the superior performance of this network.

Table 4. Difference in boundary position between the different network architectures and the manual observer for the entire data set (120 B-scans) with differing numbers of classes used during the training. The results are reported in mean values and (standard deviation) in pixel units.

CIFAR based CNN architecture					
		33x33 patch		65x65 patch	
	Layer	Mean error	Mean abs. error	Mean error	Mean abs. error
8 classes	ILM	0.05 (2.64)	0.85 (2.50)	-0.09 (1.47)	0.82 (1.23)
	INL/IPL	-0.81 (3.94)	1.52 (3.72)	-0.24 (2.16)	1.26 (1.78)
	RPE	0.10 (2.50)	0.98 (2.31)	-0.42 (3.64)	1.05 (3.51)
4 classes	ILM	-0.18 (3.83)	1.09 (3.67)	-0.25 (2.52)	0.91 (2.36)
	INL/IPL	-1.05 (6.89)	3.52 (6.02)	-0.48 (2.95)	1.64 (2.51)
	RPE	-0.00 (2.12)	1.12 (1.80)	-0.24 (1.74)	0.86 (1.53)

Complex CNN architecture					
		33x33 patch		65x65 patch	
	Layer	Mean error	Mean abs. error	Mean error	Mean abs. error
8 classes	ILM	-0.09 (0.96)	0.58 (0.77)	-0.10 (1.49)	0.59 (1.37)
	INL/IPL	-0.06 (1.41)	0.68 (1.23)	-0.03 (1.35)	0.69 (1.17)
	RPE	-0.08 (0.88)	0.57 (0.68)	-0.13 (1.43)	0.59 (1.31)
4 classes	ILM	-0.19 (2.00)	0.65 (1.90)	-0.17 (2.06)	0.65 (1.96)
	INL/IPL	0.04 (2.49)	0.87 (2.33)	-0.02 (2.17)	0.82 (2.01)
	RPE	-0.08 (1.18)	0.58 (1.03)	-0.15 (1.47)	0.58 (1.35)

3.4 Other results

Tables 2 and 3 illustrate the close agreement between the automatic and manual methods as well as the repeatability of the outcomes, and generally demonstrate only small errors between the automated and manual (truth) methods on average. However, it is also important to understand the nature of the disagreement between the methods. For all seven boundaries, the GCL/NFL generally showed the greatest differences especially near the edges of the OCT

images in the presence of retinal blood vessels as shown in Fig. 4. To better understand this, a few examples of interest are discussed here and presented in Fig. 8. In Fig. 8, the results from the manual observer (red line) are compared with the automatic method (yellow dotted line) for the GCL/NFL boundary. For each example, a close-up region where the error is located is presented at the bottom of the image, along with the probability map.

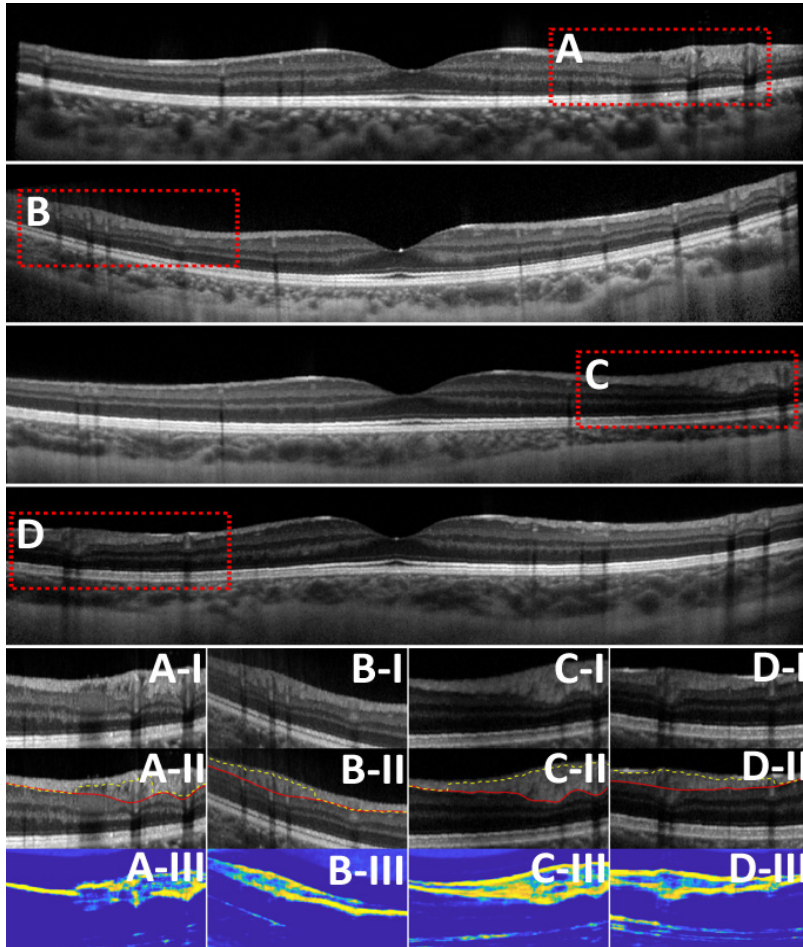


Fig. 8. Example of four B-scans with disagreement between the manual (red line) and automatic methods (yellow line), due to the presence of retinal blood vessels within the layer in these locations and the associated changes in the intensity of the boundary. The bottom panel presents a close-up of the region of interest, with the corresponding region of the B-scan (I), the manual and automatic segmentation (II) and probability map for the GCL/NFL (III).

For a number of these examples, the probability maps highlight two potential paths, above and below the GCL/NFL, given that the GCL/NFL has such large topographical variations (features) across the image, the “learned features” vary significantly. For these particular examples, the paths with the false positive coincide with the regions of the NFL where there are retinal blood vessels leading to shadows and hence reduced image contrast in the region. In other examples shown, the ILM boundary provides a stronger feature that is traced by the graph-search, similar to the way the GCL/NFL features appear close to the foveal center.

4. Conclusions

This study examined the effect of patch size and CNN architecture design to provide a novel approach for the segmentation of the retinal boundaries in OCT images. The technique builds

upon and extends a previously proposed method that combines CNN methods with graph-search [34]. The CNN method is used to detect the probability of the locations of the boundaries present in the OCT image. The boundary probability maps are then traced using a graph-search method to extract the position of the boundaries. In this work, a number of different CNN design aspects were proposed and their performance compared, demonstrating that CNN architecture has a significant impact on the segmentation results. Specifically, transitioning to a large input patch for the CNN seems to take advantage of the rich features and improved the performance and the detection of the retinal boundaries, this improvement was particular evident in the reduction of the standard deviation error. Although no previous studies have systematically examined the influence of patch size on CNN performance for the analysis of OCT images, our findings are in agreement with previous research that has shown that a larger image patch in CNNs can lead to better classification accuracy [49–51]. However, it is important to note that part of this improvement may be due to the fact that a more complex network is needed to handle the change in patch size. In this work, the effect of a more complex network architecture was also evaluated, which showed a significant improvement across all layers, in comparison to the original network. Although the improvement was more significant for the larger Complex network, even the smaller patch size Complex outperformed the two tested CIFAR networks. Overall, from the results obtained on a data set containing 120 B-scans, the proposed method shows close agreement with manual analysis performed by an experienced observer.

In this study, it was also demonstrated that the performance of this architecture depends on the number of classes used during training. Similar features across the image are likely to trigger the incorrect detection of the boundaries (i.e. false positives), which will affect the graph-search and result in segmentation error. The proposed larger and more complex network showed less dependency on the number of classes and should be considered for OCT retinal boundary classification problems with fewer classes (pre-segmented retinal boundaries).

Despite the encouraging results given by the technique in this data set of images, there are a number of potential limitations that should be considered. All the B-scans had a quality index (QI) value greater than 20 dB (the mean QI from all measurements was 33 ± 3 dB) as per the manufacturer recommendations. This value is achieved through the averaging of multiple B-scans to reduce speckle noise and enhance contrast, however, B-scan averaging may not be feasible with dense volumetric scanning protocols. To ensure the method can deal with images of lower quality, a number of options should be explored in the future. Training the CNN with labelled retinal images obtained under lower QI (fewer averaged images) or data augmentation. This second approach would involve adding noise to the data, using an OCT speckle noise model [52, 53]. It is also acknowledged that none of the subjects from the data set had retinal pathologies. Although the small square network (33x33), which is used in this work as a benchmark, has already shown good results for non-exudative age-related macular degeneration [34], the use of the proposed CNN architectures and their benefit in subjects with different posterior segment pathologies should be validated in the future.

Manual analysis of retinal boundaries is a complicated and time-consuming task. The method proposed here provides encouraging results for the automated segmentation of optical coherence tomograms of the normal human retina, with a close agreement with the results from an experienced human observer. The work presented here may assist in the future development of CNN methods for retinal boundary detection.

Funding

Rebecca L. Cooper 2018 Project Grant (DAC); Telethon – Perth Children’s Hospital Research Fund (DAC).

Acknowledgment

The Titan Xp used for this research was donated by the NVIDIA Corporation.

Conflict of Interest

The authors declare that there are no conflicts of interest related to this article.